① clean Data ⎰ missing value.
                ⎱ outlier

② feature selection : backward / forward / stepwise.

⇒ data quality record

③ Modelling

i) choose hyper parameter (eg: deep of tree).

2) Divide training set and test set

Adjust ratio , eg: 7:3   ⇒ 9:1 may overfitting

Method:

Logistic Regression

Naive Bayes

K - Nearest Neighbors   choose 3 of 6

Decision Tree

Support Vector Machines.

◦ Advanced Ensemble classifiers

◦ purpose : classification.